

# Jakub Podolak

✉ jakub.podolak.43@gmail.com • 🌐 linkedin.com/in/jakub-podolak  
podolakjakub.github.io/

## Experience

### Data Scientist

*dotData, remote*

**Warsaw (Remotely)**

*Jul 2023 – Current*

- Key decision-making position in research and development of the automated **feature discovery** platform.
- Implemented **gradient boosting** inspired methods for automated feature discovery. Prototyped new features with **OpenAI models**, utilising **vector embeddings** and **in-context learning** techniques.
- Delivered 3 new prototypical features to product, currently in trial, with one already presented to public.

### ML Research Assistant

*Ideas NCBR, Poland*

**Warsaw**

*Apr 2023 – Jul 2023*

- Implemented **early exit methods** in Computer Vision models in **PyTorch**, based on research papers.
- Benchmarked efficiency of **Dynamic Architectures** in ResNets and **Vision Transformers**.
- Performed experiments on a distributed GPU cluster with slurm workloader, with focus on **accuracy-performance tradeoff** in Machine Learning.
- Internship continuing to publication at conference.

### Intern Researcher

*University of Warsaw, Poland*

**Warsaw**

*Oct 2022 – Dec 2023*

- Researched usage of **open-source LLMs** and **OpenAI Models** to automatically detect disinformation and hate speech, and provide fact-checked responses.
- Statistical results testing - **A/B testing**, **bootstrapping** and statistical significance measuring.
- Utilised **Huggingface transformers** library to quickly prototype and evaluate models.
- Resulted in **2 articles currently in review** for the ACM Web Conference 2024.

### Junior Machine Learning Specialist

*MIM-Solutions, Poland*

**Warsaw**

*Apr 2021 – Jan 2023*

- Developed **LGBM Model** to predict the conversion rate of web ads, with the use of **Robust Neural Networks** and **PyTorch** to extract ads' image features. Using those predictions, we **increased profit per click by 90%**.
- Trained, validated and measured performance of **SVC models** for malicious website classification. Contributed to text-preprocessing pipeline, and data gathering.
- Created a computer vision pipeline for low-quality medical documents digitization with use of **OpenCV**, **custom Hough transform** and **Google OCR**.

### Intern/Junior React Developer

*Spinney, Poland*

**Krakow**

*Jun 2020 – Nov 2020*

- Developed **React.js/react-native** components and implemented updates in mobile/web game with over **5 million downloads**.

### Robotics and Python Teacher

*Primary and High Schools, Poland*

**Krakow**

*Jun 2018 – March 2021*

- Taught High School students programming in **Python** and created together a computer game with the PyGame library.

## Education

### MSc Artificial Intelligence

*University Of Amsterdam, Netherlands*

**Amsterdam**

*Sep 2023 – Jun 2025*

- Classic **Machine Learning** course - **Bayesian models**, SVM, Optimizers and neural networks.
- **Computer Vision**, **Deep Learning** courses - **Numpy** and **pyTorch**.

### BSc Computer Science

*University Of Warsaw, Poland*

**Warsaw**

*Oct 2020 – Jun 2023*

- Graduated from Algorithmics and Data Structures in C, C++ **top of class**.
- PostgreSQL, **OOP in Java**, OCaml, **Statistical Data Analysis**.
- **Deep learning, NLP course** - Gradient Descent and optimizers, **Computer Vision** and other DNN techniques .
- **Microservices course** - Building a distributed stock simulation tool with **Docker, Kafka Streams, RabbitMQ**.
- **English - C1 level** (CAE Cambridge Assessment), Polish - native.

## Organisations

---

### Member and Co-Leader of the ML in PL conference

**Warsaw**

*ML in PL, Poland*

*Dec 2022 – Present*

- Coordinating work of the volunteers and other organizers to create the next edition of one of the largest ML conferences in Central-Eastern Europe.
- Reached out to **Machine Learning researchers** and invited them to panel talks during **ML in PL** conference. Took care of the guests and kept them updated.
- Prepared questions for a panel talk and helped with onsite organization

### SocialAI

**Warsaw**

*SocialAI, Poland*

*Apr 2022 – Jun 2023*

- Utilised pre-trained polish **BERT** model to predict text sentiment in psychological research. Used **Markov Chains** and **convolution** to model emotion changes in time.

## Recognitions

---

- University representant at **CERC 2022 - Central European ICPC Contest** - biggest team algorithmic competition in this EU region.
- **1st Place in Goldman Sachs 2022 Hackathon** in Warsaw.
- 2nd place in BEST Hacking League Hackathon - **created the AI chatbot app with custom knowledge** based on OpenAI API.
- Finalist of **Polish Olympiad in Informatics**.
- 3rd place in Polish CanSat competition where we built and coded our can-sized probe. Organized by ESA (European Space Agency).
- Multiple Codeforces competitions and algorithmic camps.

## Courses and participations

---

- **NVIDIA** Fundamentals of Deep Learning certificate.
- Speaker at **Data Science Summit ML Edition 2021** - talked about robust neural networks and feature extraction.
- Speaker at **PyKonik** local python meetup in Krakow.
- **MIT Intro to Deep Learning** - MIT 6.S191 online course containing **RNNs, LSTMs, GANs, RL**.
- Participated in **EuroPython 2019/2022** in Basel and Dublin - **NLP, ML and DS techniques**.
- **Scrum/Agile** course in Ringier Axel Springer Poland.
- Speaker at Seminar in MIM Solutions company - told about **Meaningful latent image representations** using **ResNet** classifier.

## Hobbies

---

- Producing electronic music - some releases on Spotify, cooking, 3D-printing